1. **Why's of Performance Criteria**
2. **Generating FRM-like Obs**
3. **Regionality**

CASAC
Subcommittee on Particle Monitoring
January 28, 2002

EPA/OAQPS/EMAD

Why Collect FRM-like PM2.5 Mass?

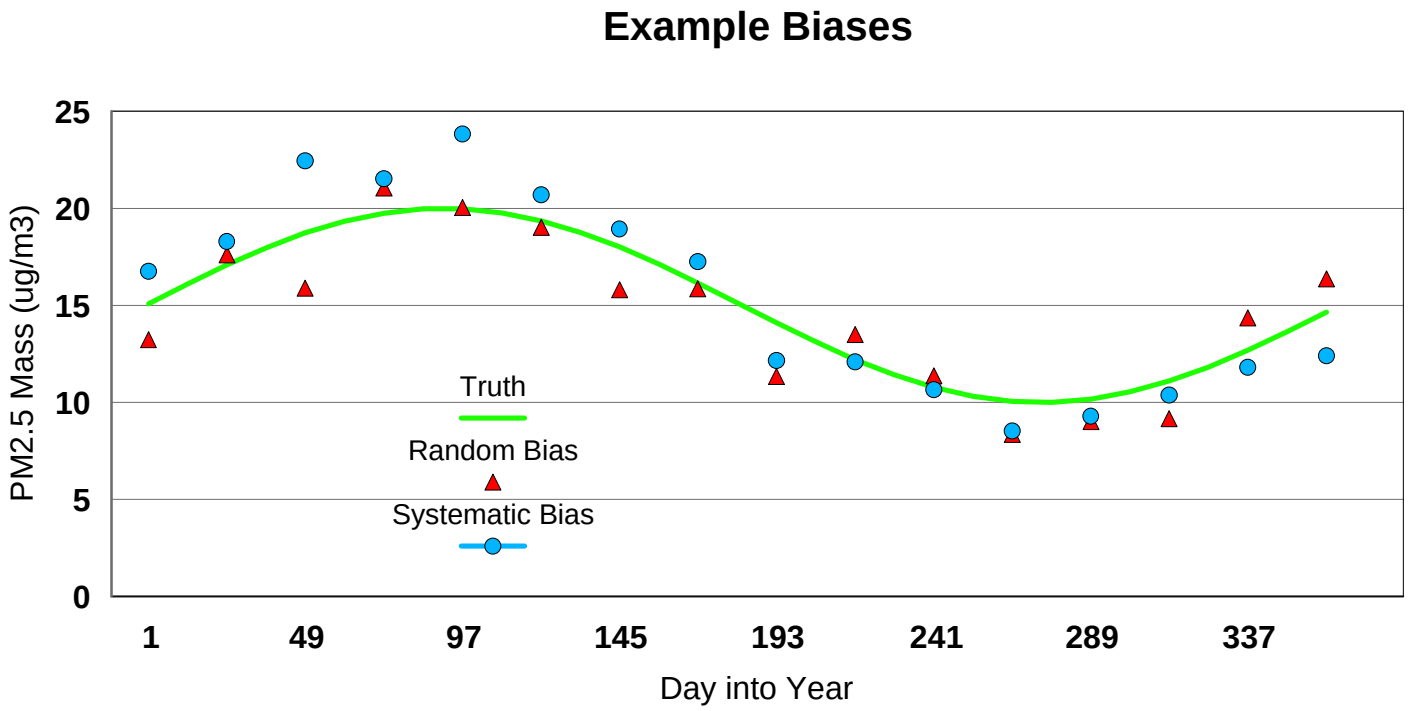
- **Consistency / comparability across space and time**
  - **Comparison with annual PM2.5 NAAQS**
  - **Comparison with daily PM2.5 NAAQS**
  - **Information for sensitive groups (AQI)**
  - **General information to public (mapping)**
  - **Support health studies, evaluation of emission inventories, simulation models, ...**
  - **General understanding / characterization (temporal and spatial) of air quality**

Data can be used for all these analyses… BUT…

real question is how confident are we in the results?

Terminology - Definition of Bias

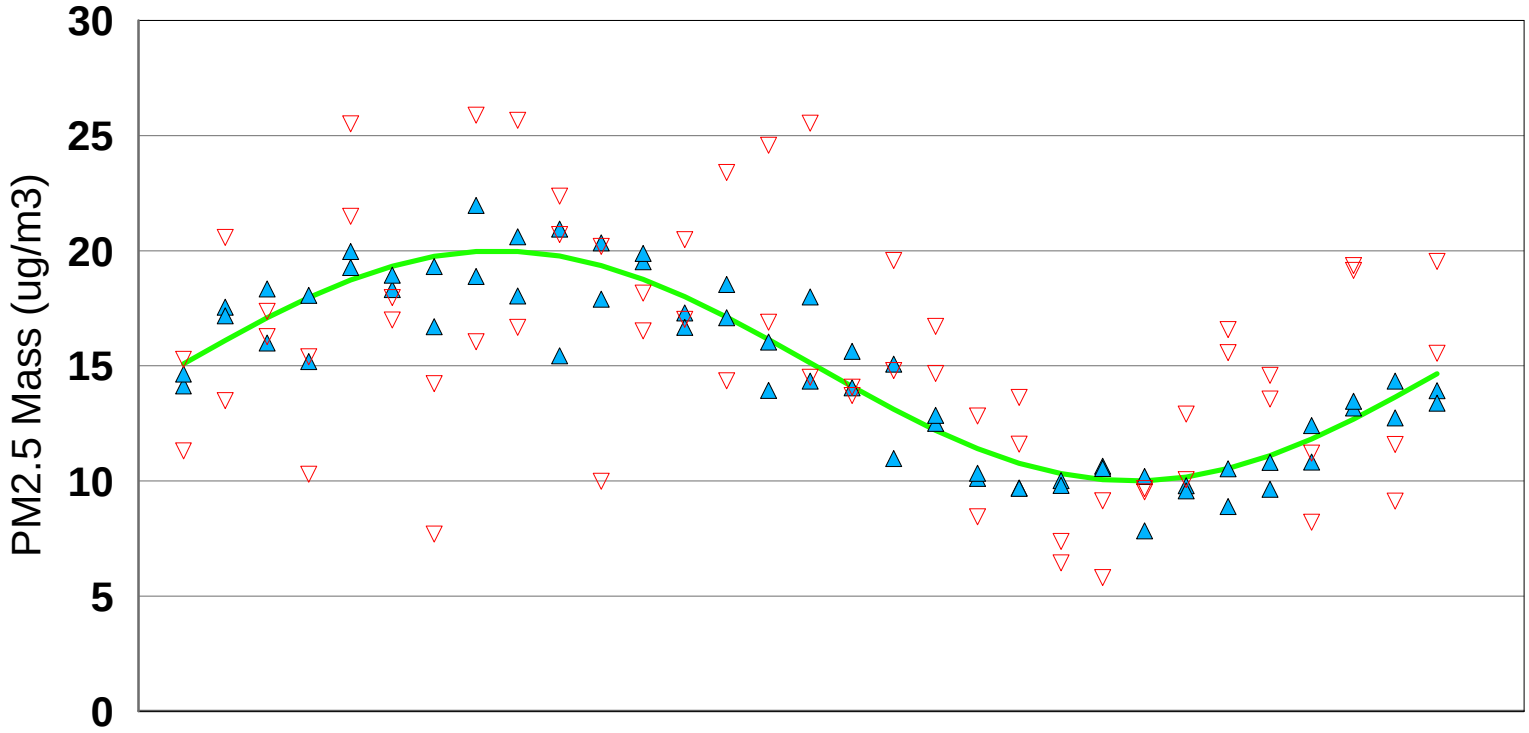
- **Bias - deviation from "truth."**
- **Estimated using PEP. Average (FRM-PEP)/PEP.**



Annual biases:	-4.0% (random)	0.5% (systematic)
Hi/Low biases:	-5%/-4% (random)	10%/-9% (systematic)

Terminology - Definition of Precision

- **Precision - repeatability of an instrument.**
- **Estimated using collocated instruments of same make. RMS of (Colo-Prim)/(sqrt(2)*avg(Colo,Prim)).**



Annual Precision: 9% (precise system) and 23% (imprecise system)

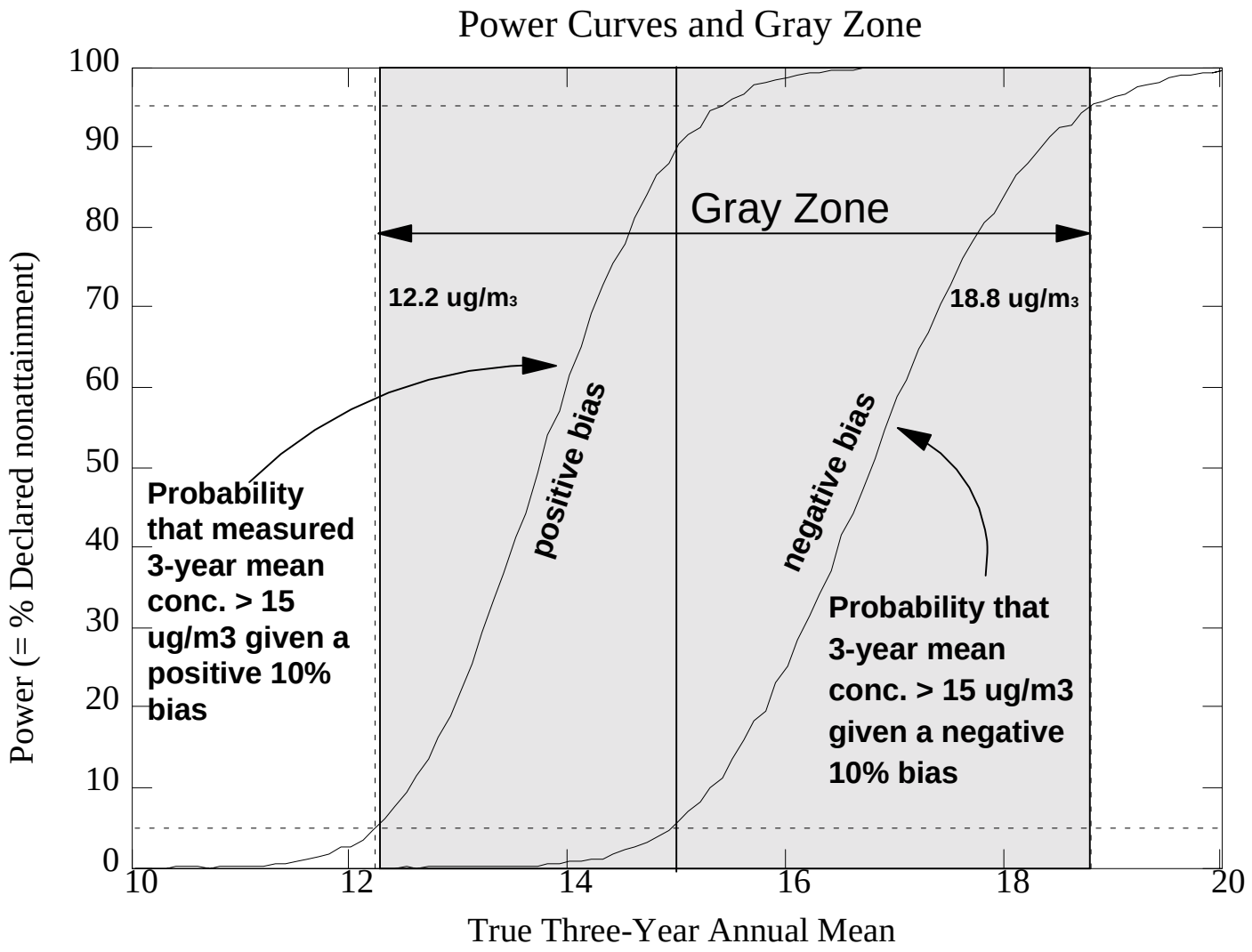
Instrument-Based
Sources of Error in Various Decisions

- To date, concerned with estimating 3-yr avg. annual conc, at a site
  - estimate impacted by overall average bias
  - not impacted by temporal pattern in bias or by precision (since averaging so many values)
- Now concerned with daily conc and public reporting, at a site
  - concerned with estimating bias for each observation and high percentile (98th) in particular
  - estimate impacted by temporal pattern in bias and confidence impacted by precision
- ... and soon to be concerned with Spatial Fields, daily / seasonal / annual surfaces of concentrations (mapping)
  - estimate impacted by spatial pattern in bias

Main Assumptions for Power Curves for Annual NAAQS

1. Annual NAAQS is controlling standard
2. Hence, bias and measurement variability (precision) applies to entire 3 years
3. 1 in 6 sampling with 75% completeness (144 days)
4. Lognormal distribution for population variability, 80% CV
5. Seasonal ratio (ratio of avg conc for highest season to lowest season) = 5.3
6. No auto correlation in daily concentrations
7. Type I and type II decision errors set to 5%

Power Curves for Annual NAAQS



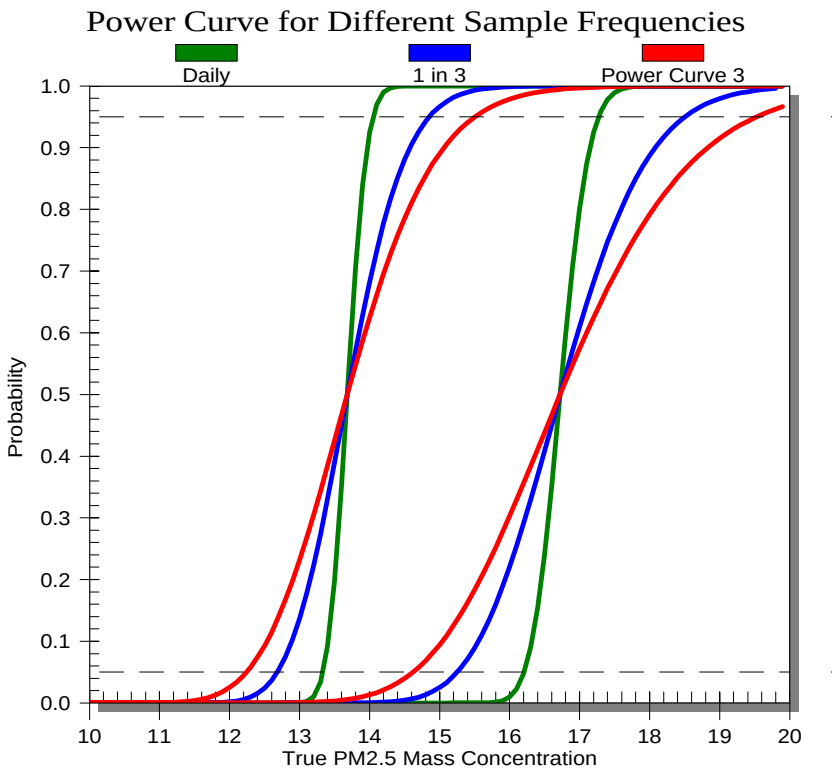
Resulting DQOs for Annual NAAQS

- **Acceptable/achievable 3-yr average bias was 10% and 3-yr measurement precision was 10% CV.**
- **Associated gray zone is [12.2,18.8]. Recall this**
  - **is for comparison to annual NAAQS, and**
  - **is for one of the most extreme cases**
    - high seasonal ratio
    - high popn cv
    - 1-in-6 sampling with 75% completeness
- **Annual Standard Gray Zone**
  - **especially sensitive to: sampling frequency, bias, population variability, seasonal ratio**
  - **not sensitive to: measurement precision**

Examples of Sensitivity of Gray Zone

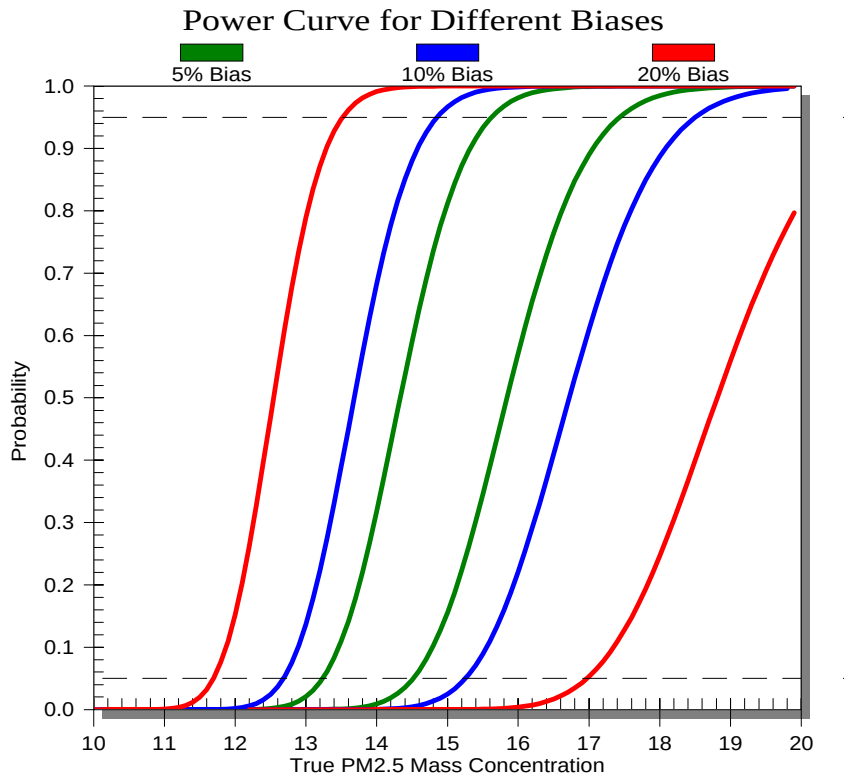
Sampling Frequency

- 1 in 6: [12.2,18.8]
- 1 in 3: [12.8,17.9]
- Daily: [13.5,17.1]



Bias

- 5% bias: [13.0,17.7]
- 10% bias: [12.2,18.8]
- 20% bias: [11.3,21.1]



Important Gray Zones

- Annual Standard assuming 1 in 6 sampling: [12.2,18.8]
- Annual Standard assuming 1 in 3 sampling: [12.8,17.9]
- Daily Standard: under development
- AQI cutpoints (~15, 40, 65, 150 ug/m^3): under development

Performance Standards for Decision Regarding Annual NAAQS

	Initial Testing for Acceptance of Method	Ongoing Performance Evaluation
REM	+/- 10% bias 20% CV (requirement)	Obtain 1-in-3 DQO gray zone [12.8,17.9] OR +/- 10% bias, 20% CV
CAC	+/- 10% bias 20% CV (goal)	Obtain 1-in-6 DQO gray zone [12.2,18.8] OR +/- 10% bias, 20% CV

For decisions involving high percentile (daily standard) or individual obs. (public reporting, mapping, spatial fields), will need to add perf. stand. for bias for each season or for quartiles of concentrations.

Perf. Standards - Evaluation

- **Continuous and FRM instruments operated by States/locals for initial testing as well as ongoing evaluation**
- **Require continuous collocated with continuous to evaluate precision (~ 15% of sites)**
- **Require continuous collocated with FRM to evaluate bias**
  - **REM: 30% of continuous REM network**
  - **CAC:**
    - 100% of continuous CAC network being used to reduce FRM sampling frequency
    - 10% of continuous CAC network for other monitoring objectives
- **Thus should have sufficient data to test and evaluate performance of continuous instruments.**

Transformations

- **Recall, overarching goal is to have comparable data.**
  - **If continuous data meet DQOs as they are (ie, continuous producing FRM-like measurements), no transformation required.**
  - **If continuous data do NOT meet DQOs,**
    - first, be sure continuous instrument operated according to good engineering principles (eg, achieving good cutpoint)
    - then, develop transformation to create FRM-like measurements.
- **For REM, only simple transformations allowed.**
- **For CAC, any transformation allowed.**
- **Transforms based on**
  - **at least 100 pairs of valid data per site**
  - **all seasons being well represented, not necessarily in one year**
- **Transforms apply to data collected in the future, not the past**

REM Transformations

- **For REM, only "simple" transformations will be allowed**
  - **Goals are to predict accurately FRM concentrations *and* for two independent analysts to develop identical transform**
  - **"Simple" means FRM = intercept + (slope*Continuous)**
  - **Lots of details**
    - number of required pairs (~100 spread evenly through year)
    - log transform before fitting intercept and slope?
    - identification / handling of influential pairs
    - handling of negative or zero concentrations
    - range of pairs (limit inference to range?)
    - way to estimate slope and intercept
    - way to determine if transform good enough to be used (R^2, MSE, meet DQO)

CAC Transformations

- **For CAC, transformation flexible**
  - **Goal is to predict accurately FRM concentrations**
  - **FRM is response variable, continuous is an explanatory variable, other explanatory variables allowed**
  - **Transform may be nonlinear in the explanatory variables**
  - **How to determine if transform good enough to be used (R^2, MSE, meet DQO)**
    - Possibility: Transform considered good enough to be used if number of pairs is > 100 and resulting R^2 is > 0.80 (based on DQO for AQI reporting).

Transformations - Additional Issues

- What is spatial representativeness of transformation?
- What is temporal representativeness of transformation?
- Frequency transformation checked?
- When transformation no longer sufficient to meet performance specs,
  - what is done with previously transformed data?
  - how to transition from one transform to the next?
- What goes into AIRS?
  - FRM-like measurements?
  - Field indicating form of transform so transform can be "undone" if raw data are of interest?

Regionality of Transformations

- **Could have site-specific transformations. Implies testing and ongoing evaluation would have to occur at each site.**
- **Instead, find regions of country within which one transformation can be used to produce FRM-like measurements.**
- **Determine "Regions" a priori**
  - **divvy country into regions where continuous/FRM relationship should be similar**
  - **regions could be decided based on**
    - principles about instruments, PM2.5 composition, meteorology, …
    - observed relationships
- **Let the data determine "Regions"**
  - **look at collocated FRM/continuous sites**
  - **see where relationship is consistent, hence no transform or the same transform can be used**
  - **regions may be any size/shape and may vary by method**

Regionality - Issues

- How to determine regions? A priori or data driven?
- Several details about developing one regional transformation based on data from several collocated sites
  - pool data to derive estimates or pool estimates?
  - weight sites based on # pairs?
  - use transformation from site with greatest range?
- How frequently are region definitions reevaluated?
- If/when regions change through time, how does this affect use of previously-approved continuous instruments?